

Route planning on orienteering maps with least-cost path analysis

Gáspár Albert ^{a,*}, Zsófia Sárközy ^b

^a Eötvös Loránd University ELTE, Institute of Cartography and Geoinformatics, albert@ludens.elte.hu

^b Eötvös Loránd University ELTE

* Corresponding author

Abstract: The feature categories of an orienteering map are prepared to allow the map reader to estimate the travel time between any two points on the map with a good approximation. This requires not only an accurate map, but also a key that adapts to the speed of travel. Such map key is developed and maintained by the IOF (International Orienteering Federation), and technically all the orienteering maps are compiled by using it. Estimated time also plays an important role in planning the courses of orienteering races. The course setter estimates time based on a route he thinks is ideal, but the speed of travel is basically a non-linear function of terrain, road network and land cover. Because of this, the easiest (ideal) route between the two points and its time cost can be calculated using the least-cost path (LCP) GIS method, which can be prepared to take into account these three map feature categories. This method is based on the calculation of a cost surface, then the analysis of the ideal path from a given point to the destination. The automation can be adapted to any orienteering map due to the similarities of the map keys. This study shows that if the weight corresponding to the different feature categories is given properly, the ideal path between two points on orienteering maps can be calculated. The ideal path, however is still a subjective category, which may depend on the capabilities and preferences of the orienteer. In this study the routes calculated with the LCP method were compared with the suggestions of the ideal routes by orienteering runners of different ages. The results show that the routes given by sportsmen can be simulated with the LCP method and even the time cost of the calculated routes can be calculated. This study can lay the groundwork for a GIS tool helping the course setting process on standard orienteering maps.

Keywords: least-cost path, orienteering, route planning, raster map

1. Introduction

Route planning algorithms are frequently used in modern navigation tools, including smartphones and help to find the most practical way between a departure and a destination point. The algorithms are based on the mathematical analysis of the existing road network taking into account different attributes of the road sections to precisely estimate the time-, and distance-costs of the route (Luxen & Vetter, 2011). Navigation tools such as Google Maps require a digital road map, which can be considered mathematically as a network, and the analysis concentrates on finding the shortest path that connects key nodes in the network (Delling et al., 2009). On those places where road network is sparse, yet some sort of traffic is existent, the route analysis requires a different data structure, which is the raster model (Herzog, 2013). Since such traffic is not usual, and more importantly depends on the participant, route planning methods developed for raster maps are not included in well-known navigation applications.

Maps in raster data structure may contain any kind of surface type including roads, but in a similar way as the different cover-types: a regular grid of points contains the information. The denser the points are, the smaller the area that is represented by one point is, and also the more precise the raster-type map will be. The most common regular grid is the square grid, and it can also be considered as a network where a certain node has usually eight

neighbours (Mitchel, 1988; van Etten, 2017). Since the physical distance between the grid nodes are constant, the route planning algorithms in raster maps analyses the differences of attribute values assigned to the node points (Mitchel, 1988). The raster map used for the analysis thus contains numerical values representing the “travel cost” of a given surface type, and is therefore called a “cost raster” (e.g. Herzog, 2013; Alberti, 2019).

The cost raster can be processed to calculate the travel cost to a destination point. The closer the departure point is to the destination, the smaller the travel cost will be, so the route planning algorithm calculates the cost cumulatively from the destination point, and creates a “cost accumulation raster” (Etherington & Holland, 2013). The ideal route between a departure point and the destination point is planned on the accumulation raster, by selecting consequently the grid node in the neighbourhood, which has the smallest accumulated cost value (Dijkstra, 1959). The algorithm is called the least-cost path analysis (Douglas, 1994; Alberti, 2019).

The least-cost path (LCP) method is usually appropriate for analysing wildlife migration, human travel in archaic times, or planning trace-lines of yet to be built roads (Herzog, 2013; Alberti, 2019). In the case of traveling on foot at hiking speed, even the traveling time can be estimated with Tobler’s off-path function (Tobler, 1993). This function works with the slope degree, and does not take the land cover into account. Orienteering maps,

however are prepared for runners, who solo navigate on an unmarked course searching for control points in a strict sequence, and the hiking speed function is useless in their case. The courses on an orienteering event only partially rely on using trails or roads, and in most cases the control points are physically placed on landmarks, where trails do not lead to. Also, the speed of an orienteering runner depends not only on the slope, but the roughness of the terrain and especially the vegetation type that covers the crossed area.

We assume that in the case of orienteering maps the LCP method can help to calculate the expected finishing time of an orienteering race, which is currently estimated by the course setting staff trying to meet the requirements of the International Orienteering Federation's (IOF) rules (Hébert-Losier et al., 2015). The main complication in applying the LCP occurs during the calculation of the cost raster, since not only the hypsometry and roads, but the coverage types are also taken into account. The LCP should be executed on an accumulation surface, which is carefully calculated from all these different map features (Herzog, 2013). The present study aims to demonstrate how such a cost raster can be calculated from an orienteering map for applying the LCP on it, and with using empirical approaches even the travel time can be estimated.

2. Materials and methods

The test area is a well-known and frequently visited recreational forest in the western mountainous suburban area of Budapest, Hungary (fig. 1). There are several versions of the map, maintained by the orienteering clubs' responsible map compiler staff in the Central Hungary region. The currently used version¹ of the map (Dénes et al., 2014) was last updated in OCAD² software in 2019 and covers 5.17 square kilometres. The original OCAD file was converted into various shp-files differentiated by the type and the geometry of the features. Shape files can be processed in any GIS environment and can contain point-, line-, and area features with attributes. In the present study the QGIS 3.14 and SAGA GIS open access programs were used. From the OCAD-file, the contour lines and the road network were extracted as lines, and the coverage types as area features into separate data layers of the GIS. Point-like objects were processed only if they played important role in the preparation of the cost rasters.

The processing started with the preparation of the data layers, then we selected start and finish points on the map in order to carry out the least-cost route model testing. These routes were also tested with two groups of orienteers: professionals and average. Finally, the model parameters were adjusted until they best fitted the route of the professional orienteers and then compared with the route of the average runners.

2.1 Preparation of the map data layers

The map data layers (MDL) are vector-type geodatabases (shapes) which are used to create the cost rasters of the hypsometry, road network and the coverage. Each of the MDL-s require a unique approach to prepare them for further processing.

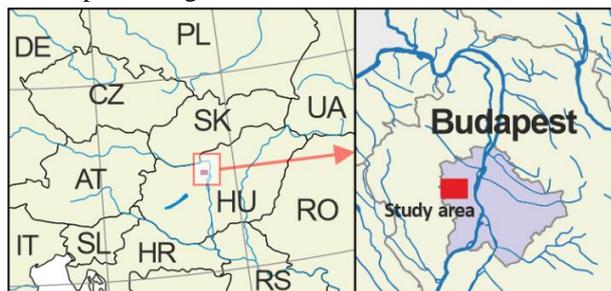


Figure 1. Location of the study area.

The hypsometry MDL in its final state contains the contour lines and some height and depth points. Since lidar survey or fine-resolution digital terrain model (DTM) do not exist in the area yet, the contour lines of the map were used as base data for creating a 1 m-resolution DTM. The contours were tagged with the elevation data, and the node points of each contour line were extracted into a point database which was used to create the DTM. Since this method can produce false morphology, namely "terraces" between the data-poor inter-contour areas, the kriging interpolation was used with exponential variogram model and a relatively large (40 m) lag distance, which was further smoothed with a 7 by 7 m kernel of the moving average filter iteratively 9 times. The result – although the exact elevation values were changed due to the filtering – was not dominated by the false terrace-morphology marks, but kept the morphological details and – more importantly – was appropriate to calculate the morphometric features such as the slope angle (fig. 2).



Figure 2. Hill shading applied on the DTM prepared from the contour lines demonstrating the smoothness of the modelled surface on the "inter-contour" areas.

The road network MDL contained not only the roads and trails but all the linear features that may play some role in the navigation during the orienteering events (e.g. earth banks, gullies, watercourses, etc.). The geodatabase contained weights/costs (0.1 to 1) of the linear elements

¹ Provided by the Budapest Orienteering Association

² Special map-editing software for orienteering runners (<https://www.ocad.com/en/>)

depending on the relative effort used by a runner to proceed on, or over them. The impassable elements got a weight of 10 (tab. 1). The minimum value was set to 0.1 instead of 0 in order to get non-zero accumulation costs.

The coverage MDL contained all the area features of the map. Similarly to the linear features, the weights (costs) of the components were determined on the basis of the IOF map keys (IOF Map Commission, 2019). The range was normalized between 0.1 and 1, and the impassable features were set to 10.

Type	Feature name (ISOM 2017-2)*	Weight
Linear	Earth bank, (small, trench)	0.6, (0.3, 0.5)
	Earth wall	0.3
	Erosion gully (small)	0.5, (0.3)
	Cliff	0.7
	Small crossable watercourse, narrow marsh	0.3
	Hedge	0.4
	Wide road, road, vehicle track, footpath	0.1
	Small footpath	0.15
	Less distinct small footpath	0.2
	Narrow ride or linear trace through the terrain	0.2
	Railway	0.15
Area	Open land (&trees)	0.2
	Rough open land &trees, (&bushes, &thickets)	0.25 (0.45, 0.6)
	Forest	0.3
	Vegetation: slow running (&good visibility)	0.35
	Vegetation: walk (&good visibility)	0.4 (0.45)
	Vegetation: fight	0.75
	Paved area	0.1
Impass.	Impassable fence	10
	Impassable cliff	10
	Area that shall not be entered	10
	Building	10

Table 1. Linear and area features and their weights in the cost-surface calculations. The ISOM 2017-2 is the adjusted version of the International Specification for Orienteering Maps published in 2019 (IOF Map Commission, 2019).

After preparing the MDLs a mask with a weight of 10 was applied on each of them covering the perimeters of the mapped area.

2.2 Carrying out user tests

Two different user tests were prepared to ensure the quality control in the present study. A group of experts (elite orienteers) was involved in the study to provide reference for designing the parameters of the LCP modelling, and a group of orienteers as “users” to provide material on which we could test the results. The two groups consisted of mainly different people.

In the first phase, a total of 11 coaches and experienced athletes were asked to draw the ideal path between eight pairs of points on two variations of the map: 1) containing only the roads, trails and cover-types as like as the whole area was flat; 2) containing only the contours and other hypsometry objects (fig. 3).

On each of the maps, eight single-control (start-finish) courses were set up. The finish point was the same for all of the courses due to practical reasons: this made the GIS processing quicker because the accumulation surface was calculated only once-per-model setup. The course legs were designed to be difficult “route-choice” types to maximize the difference in the users’ answers.

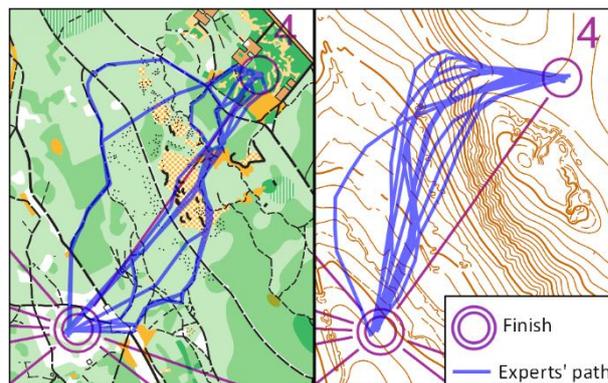


Figure 3. Excerpts of the “roads & coverage” and the “contours” test-maps showing the substantially different paths drawn by the experts on the course No. 4.

The user test was carried out with the participation of 26 junior, adult and master orienteering runners (18 women and 8 men). Their experience level ranged from beginner, with less than 2 years orienteering, to expert (more than 5 years running). Each of them got the same course setting as the first group, but in their case the map contained all the map features.

In both cases the participants drew the paths by hand and the trace-lines were digitized afterwards in the GIS environment.

2.3 Calculation of the cost surfaces and the least-cost paths

For the LCP analysis, the SAGA GIS program was used. To ensure that the machine-generated LCP is appropriate, we applied a two-step method: 1) validation of the weights used in the MDL preparation, 2) combining the cost rasters of the MDL using different weights, and matching the results with the user-tests.

The setting up of the proper weights was done empirically. The MDLs were evaluated separately and the parameters and weights were accepted when the machine-generated path followed a route that fits to the one that was suggested by orienteers from the experts’ group.

If the machine-generated path did not match with at least one (ideally the majority) of the experts’ path, the weights were re-adjusted (fig. 4).

For the road-network and coverage cost rasters the finally accepted weights are shown in table 1. At first, a cost raster for both the linear and the coverage features were

calculated with the QGIS rasterization tool, but for the subsequent steps the line network and the coverage cost rasters were patched, as if the roads and other linear elements were also area-type features having a one pixel “width”. This combined area cost-raster (ACR) was iteratively generated from the MDLs during the process and used for creating LCPs which were compared to the experts’ paths.

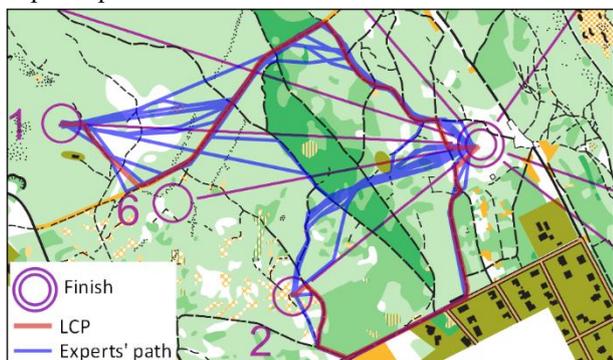


Figure 4. Excerpt of the “roads & coverage” test-map showing the paths drawn by the experts on the single-point courses No. 1 and 2, and the least-cost paths calculated by the GIS program on the ACR.

2.4 Components of the hypsometry cost raster

The hypsometric and coverage properties of an area can be very different in other study areas. It also means that the cost rasters should be calculated in a generally applicable way. In our method the generalization was achieved by data preparation based on normalisation and distribution analysis.

The hypsometry cost raster (HCR) represents morphometric data, which can be derived from the hypsometry DTM. In Tobler’s calculation, the slope steepness was the only morphometric element in the equation (Tobler, 1993). In the present study we examined the possibility to use other morphometric variables, which could be used for fine tuning the calculations. Although the slope steepness is the most important factor, non-directly it is present in other morphometric derivatives as well. The following list of variables was considered and tested during the calculations:

- Terrain Ruggedness Index (TRI) - it indicates how complex the terrain is around a cell. The complexity of a terrain refers to the standard deviation of the height around a certain location (Olaya, 2006). In the present study a 1-cell buffer was used. In the case of uphill trajectory the high TRI value refers to steep slope. In downhill trajectory the rough terrain means a higher the risk of injuries (Minetti et al, 2002);
- Plane curvature (PC) – it is useful to detect sharp ridges, spurs and side valleys, which makes the progress difficult when running in constant height. In the calculations, the absolute value of this variable was used in order to avoid negative costs;

- Slope angle (SLO) – it effects the speed of running in both uphill and downhill trajectory, the steeper the slope, the slower the running is (similarly to the TRI). The slope angle was used as a weight factor for a single cell.

Since the variables are numerically different, the normalization is necessary before combining them. Similarly to the linear and coverage MDLs, the ‘cost values’ of the variables were normalized between 0.1 and 1. The distribution function (histogram) of the variables was also taken into account: a constant multiplier c was implemented in the calculations to match the peaks of the histograms calculated from the cost raster of the linear and areal features (ACR) and the one that was calculated for the hypsometry (fig. 5). The cost raster for the hypsometry was calculated with the following equations:

$$HCR = TRI \times \sqrt{abs(PC) \times SLO} \quad (1)$$

$$c = f_{ACR}(\mu) / f_{HCR}(\mu) \quad (2)$$

$$HCR_a = c \times HCR \quad (3)$$

where the μ mark the peak values in the histograms, HCR_a is the adjusted hypsometry cost raster (HCR) value, and each variables represent cell values. The c equals 2.12 in the study area.

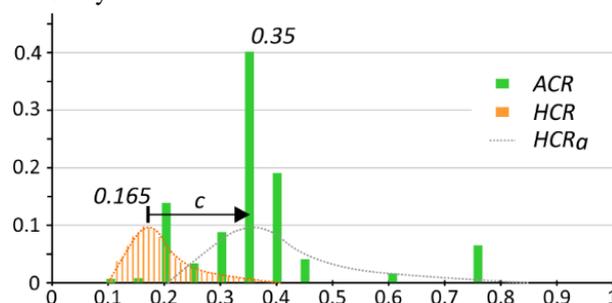


Figure 5. Diagram of distributions vs. relative proportion of cost values in the ACR and HCR. The HCR_a is the adjusted distribution of the hypsometry cost raster.

2.5 Combining the cost rasters

The preparation of the ACR and the HCR resulted two cost rasters, which were appropriate to reconstruct those paths that the experts estimated to be the bests. As base materials, these cost surfaces were used in the combination of the overall cost surface.

The two rasters had the same range (0.1-1) and due to the adjustments, the same maximum in the histogram. To carry out raster calculations, the extent, resolution (1 m) and projection also had to be the same for both of them.

The process of combination was done with raster calculations in QGIS. The two rasters were weighted proportionally: e.g. if one was weighted 0.4, the other was 0.6 so the range of the combined raster remained 0.1-1. Several variations were created and used for LCP calculations. The results were evaluated based on the user tests.

2.6 Summing up the athletes' paths

A summed athletes' path was calculated from all the digitized paths drawn by the participants. The summing was based on a heat-map, compiled from the node-points of the paths. The process was done using the "densify by interval" tool with 5 m gap. The heat maps were calculated with a 10 by 10 m ground resolution from the node points and for each of the eight courses. The reciprocal grids were then calculated from the heat-grids, and were used for calculating the accumulation surfaces. Finally the least-cost paths were also calculated on them representing the quasi "centre-lines" of the heat-maps, where originally the largest heat values were found.

The summed paths (SP) for each courses show the most often chosen routes, but in many cases there were popular second or third alternatives as well (fig. 6).

3. Results

The variations of the combined cost rasters were created to compare them with the orienteer athletes' path. The comparison was based on visual evaluation and numerical analysis. We found that the combination of the two cost rasters were the best when the HCR_a had 60% and the ACR 40% weight.

The numerical comparison included the calculations of the lengths, the total uphill, and downhill reliefs on the paths between the start points and the finish (see table inlet in fig. 6).

The machine-generated paths were almost identical to the summed paths of the test courses in the cases of course no. 3, 4, 5. In these cases both the path geometry and the physical parameters were similar (there were only a few percent difference). In the case of course no. 6 the LCP differed from the summed path but matched with popular alternative path. On courses 7 and 8 the LCPs partly matched with the summoned paths, but there were sections where they followed less popular routes (one section of the path no. 7 even traced an unmarked path).

The courses no. 1 and 2 showed the most significant differences between the geometry and the parameters of the LCPs and the SPs. However, the machine-generated path matched with some athletes' path and was much shorter than the SP, paired with only a slightly increased uphill relief on course no 2. The machine-calculated LCP of course no. 1 even contained less up-, and downhill relief than the SP.

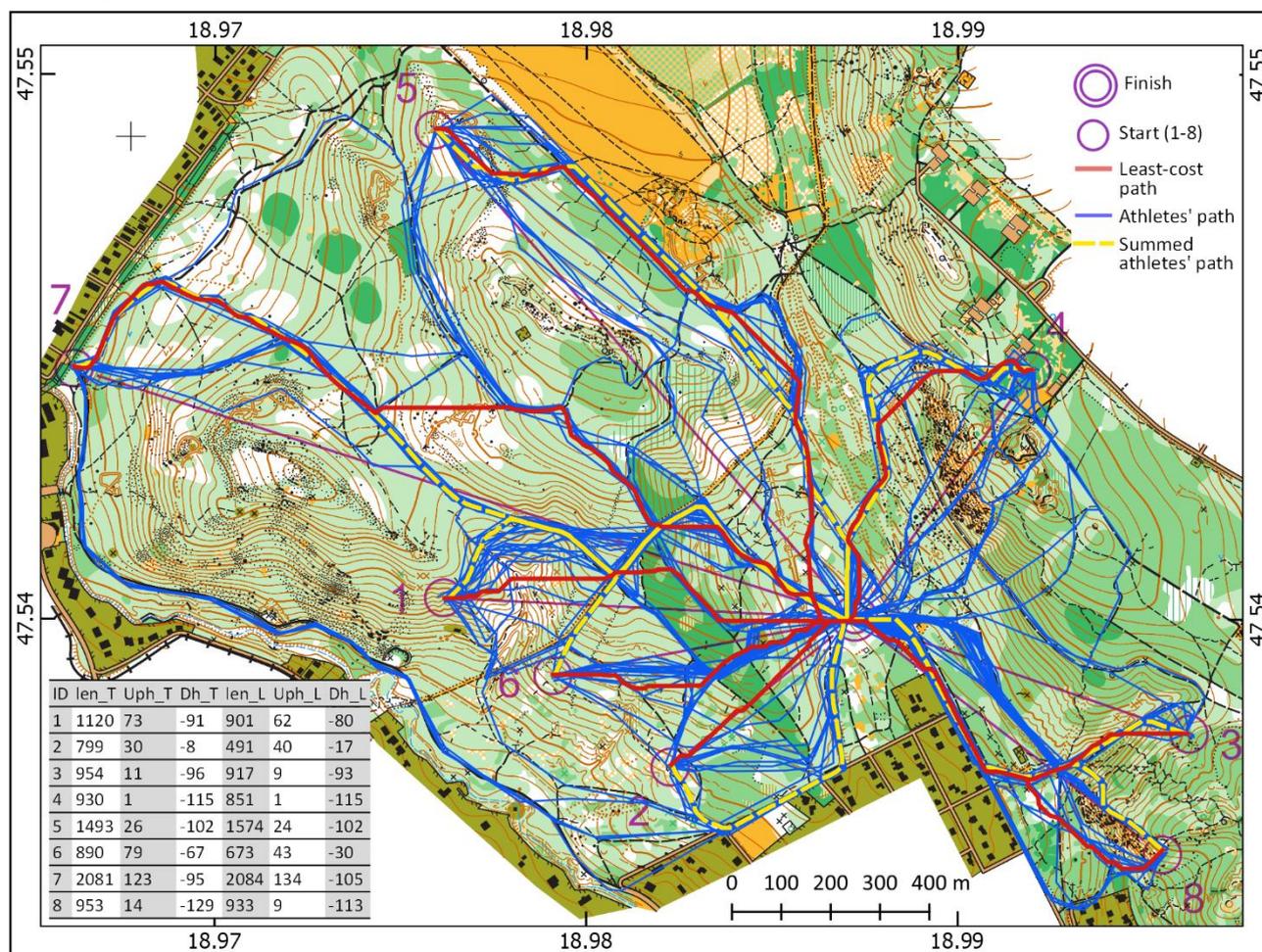


Figure 6. Map of the athletes' paths, the summed athletes' paths, and the machine-calculated least-cost paths for all the 8 test courses. The inlet table shows the length, the total uphill and downhill reliefs on the summed paths (len_T, Uph_T, Dh_T), and on the calculated least-cost paths (len_L, Uph_L, Dh_L).

4. Discussion

The LCP calculations are based on numbers and avoided the seemingly suitable paths if there was a steep slope on it. This explains why the LCP and the SP of course no. 7 differs: on the 2/3 of the summed path there is a steeper slope, which is better to avoid. However, it is very likely that taking the LCP and the SP in this case would require similar effort and time.

Evaluating a path's goodness (and more importantly its goodness over another path) requires physical testing on the field with the participation of orienteers. Such a test could be influenced by many known, and unknown factors (e.g. weather conditions, fitness and preparedness of the athletes, etc.), but it provides first-hand measurements as base data instead of data derived from spatial information (such as a map or DTM). Although, in the present study such field test were not carried out, the study area is the place of a traditional orienteering event, which is famous of its "route-choice" challenges. The courses are the same in every year and the point-setting more-or-less follows the pattern originally planned by Henrik Ripszám, founder of the Hungarian orienteering sport, in 1925 (Hegedűs, 2016). The results and the gps-tracks (if provided by the participants) were uploaded on the Hungarian Orienteering Federation's website³ and can be queried for the event in 2014. Although the records are not complete, some of the better athletes' paths were available together with the time data, and some of these paths matched with the LCPs created for the same pairs of control points (course legs) using our combined cost surface. The "better athletes" were those, who finished within 10% time plus compared to the winner.

The query aimed to estimate the computed LCPs time cost. Although, this calculation is considered as a very rough estimation, we found that one unit of the calculated cost equals approximately 1.1–1.8 seconds in the investigated cases. Applying these values on the eight pairs of points in the present study we found these estimations realistic (tab. 2).

Course	Length	Uphill	Cost	Time 1	Time 2
	m	m		min.	min.
1	901	62	345.85	6.34	10.38
2	491	40	167.23	3.07	5.02
3	917	9	282.13	5.17	8.46
4	851	1	257.50	4.72	7.73
5	1574	24	388.36	7.12	11.65
6	673	43	235.55	4.32	7.07
7	2084	134	530.45	9.72	15.91
8	933	9	306.67	5.62	9.20

Table 2. Physical parameters, calculated costs, and estimated times for the eight test courses. The "Time 1" and "Time 2"

columns represent the 1.1 and 1.8 second/cost unit calculations in decimal minutes.

Like all models, the calculated cost surface omits many of the physically existing factors. Some are irrelevant, because their effects are too small, and some of them were excluded due to the improper source data. The hypsometry MDL did not include the boulder-fields and the point-like landforms such as knolls, pits, and depressions. Also the broken-ground and stony-ground was not processed due to the inconsistency of the original OCAD file (e.g. the stony-ground's dot-symbols were digitized sometimes as area, and sometimes as point-features covered by other map objects). This type of problem can be overcome if lidar data is available.

Though the preparation of the cost rasters aimed to generalize the method with applying normalization and histogram transformation, other orienteering maps need to be tested.

5. Conclusion

The route planning on raster-type maps with various map data layers (MDLs) were successfully tested on the sample area. A new method was worked out based on the least-cost path analysis, and the raster-to-vector conversion of the orienteering maps.

The novelty of our solution is the determination of the weights of the various features, and the method of calculating the combined cost raster, which includes the hypsometry, line-network and coverage types. These types of data layers are not specific to orienteering maps, so in any case where the best route for off-road navigation is to be determined, the principle basis of the method presented here can be used.

With the here presented method and study area, the conversion of the calculated cost units into time units is possible and produces realistic results. However, the further investigation of the method is necessary and requires involving field tests with orienteers and other study areas.

Based on the results of the presented study, the least-cost analysis of complex map data layers can be implemented into standard GIS processes (in the form of plugins or applications). Such GIS tool would reduce the risk of subjective estimations in the course setting process on standard orienteering maps.

6. Acknowledgements

From the part of G. Albert the project was supported by the National Research Development and Innovation (NRDI) Fund of Hungary, financed under the Thematic Excellence Programme TKP2020-NKA-06 (National Challenges Subprogramme) funding scheme. The authors express their thanks to the athletes of the Tabán Spartacus

³ <http://rg.mtfisz.hu/rg2/#512&course=1> [last accessed on 07. 04. 2021.]

Sports and Environmental Association for participating in the user tests.

7. References

- Alberti, G. (2019). movecost: An R package for calculating accumulated slope-dependent anisotropic cost-surfaces and least-cost paths. *SoftwareX*, 10, 100331.
- Delling, D., Sanders, P., Schultes, D., & Wagner, D. (2009). Engineering route planning algorithms. In *Algorithmics of large and complex networks* (pp. 117-139). Springer, Berlin, Heidelberg.
- Dénes, Z., Sulyok, Á., Vancsik, N. (2014). Vadaskert (Hüvösvölgy). 1:10000. Budapest Orienteering Association.
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1), 269-271.
- Douglas DH (1994) Least-cost path in GIS using an accumulated cost surface and slope lines. *Cartographica* 31:37–51
- Etherington, T. R., & Holland, E. P. (2013). Least-cost path length versus accumulated-cost as connectivity measures. *Landscape Ecology*, 28(7), 1223-1229.
- Hegedűs, Á. (2016) Ripszám Henrik és a magyar tájfutás kezdetei 1925-1948. (*Henrik Ripszám and the beginnings of Hungarian orienteering 1925-1948*), only in Hungarian, ISBN 978-963-12-4521-9, 154 p.
- Herzog, I. 2013 Review of *Least Cost Analysis of Social Landscapes*. *Archaeological Case Studies* [Book], Internet Archaeology 34. <https://doi.org/10.11141/ia.34.7>
- Hébert-Losier, K., Platt, S., & Hopkins, W. G. (2015). Sources of variability in performance times at the world orienteering championships. *Medicine & Science in Sports & Exercise*, 47(7), 1523-1530.
- IOF Map Commission (2019). International Specification for Orienteering Maps ISOM 2017-2. *International Orienteering Federation*, Karlstadt, Sweden, p. 40.
- Luxen, D., & Vetter, C. (2011). Real-time routing with OpenStreetMap data. In: Proceedings of the 19th ACM SIGSPATIAL international conference on advances in geographic information systems, (pp. 513-516).
- Mitchell, J. S. (1988). An algorithmic approach to some problems in terrain navigation. *Artificial Intelligence*, 37(1-3), 171-201.
- Minetti, A. E., Moia, C., Roi, G. S., Susta, D., & Ferretti, G. (2002). Energy cost of walking and running at extreme uphill and downhill slopes. *Journal of applied physiology*, 93, 1039–1046.
- Olaya, V. (2006): Basic Land-Surface Parameters. In: Hengl, T., Reuter, H.I. [Eds.]: *Geomorphometry: Concepts, Software, Applications*. Developments in Soil Science, Elsevier, Vol.33, 141-169.
- Tobler, W. (1993). Three presentations on geographical analysis and modeling. National Center for Geographic Information and Analysis, Technical report 93-1.
- van Etten, J. (2017). R package gdistance: distances and routes on geographical grids. *Journal of Statistical Software*, 76(1), 1-21.